Considerations for Evaluation of Diagnostic Performance

Felipe A. Medeiros, M.D., Ph.D.
Professor of Ophthalmology

Hamilton Glaucoma Center
University of California, San Diego

UCSD GLAUCOMA CENTER

I have the following financial disclosures:

- **Research Support:**
  - National Eye Institute R01 EY021818
  - Alcon, Inc.
  - Allergan, Inc.
  - Carl-Zeiss Meditec, Inc.
  - Heidelberg Engineering
  - Optovue
  - Reichert
  - Sensimed
  - Merck, Inc.

UCSD GLAUCOMA CENTER

Potential Applications of Imaging Instruments

- **Diagnosis**
  - Decrease diagnostic uncertainty in those SUSPECTED of a condition (suspected of having damage or suspected of having progression)
- **Screening**
  - Identify abnormal or suspected cases in the general population or pre-selected subjects (e.g., older population, positive family history)
- **Prognosis**
  - Determine risk of developing a condition

UCSD GLAUCOMA CENTER

Potential Applications of Imaging Instruments

The design of the study should take into account the purpose of the test and involve the clinically relevant population

UCSD GLAUCOMA CENTER

Diagnostic Accuracy Studies

- Diagnostic tests are used to decrease uncertainty about presence of a condition
  - **Cross-sectional assessment:** The test is applied at a single point in time in those suspected of having the disease

Does this patient have glaucoma?

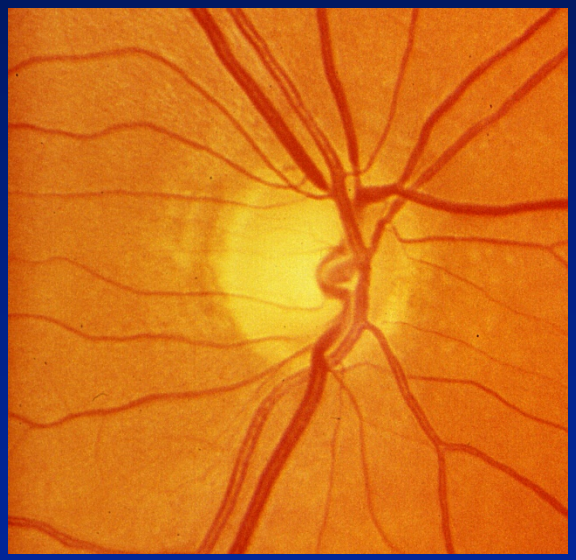
  - **Longitudinal assessment:** The test is applied multiple times during follow-up in suspects or those with confirmed disease

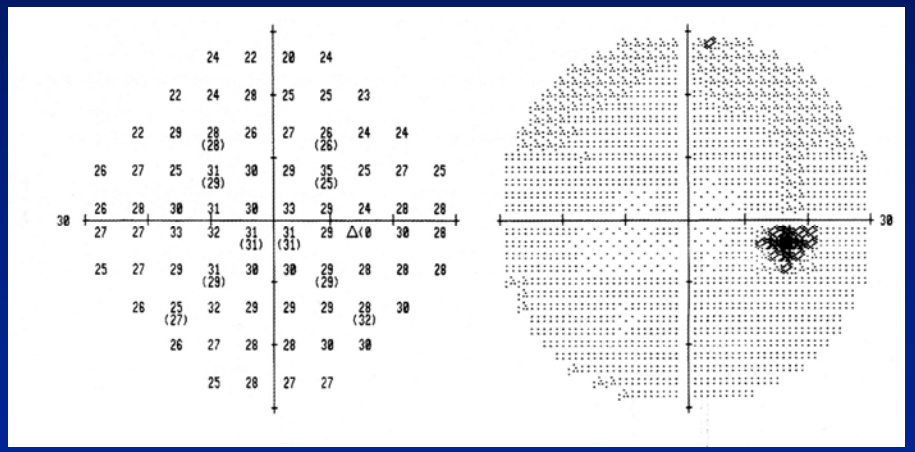
Does this patient have disease progression?

UCSD GLAUCOMA CENTER

Uncertainty is what characterizes glaucoma suspects

Suspicions optic disc appearance, with normal or suspicious visual field results





Does this patient have glaucoma?

UCSD GLAUCOMA CENTER

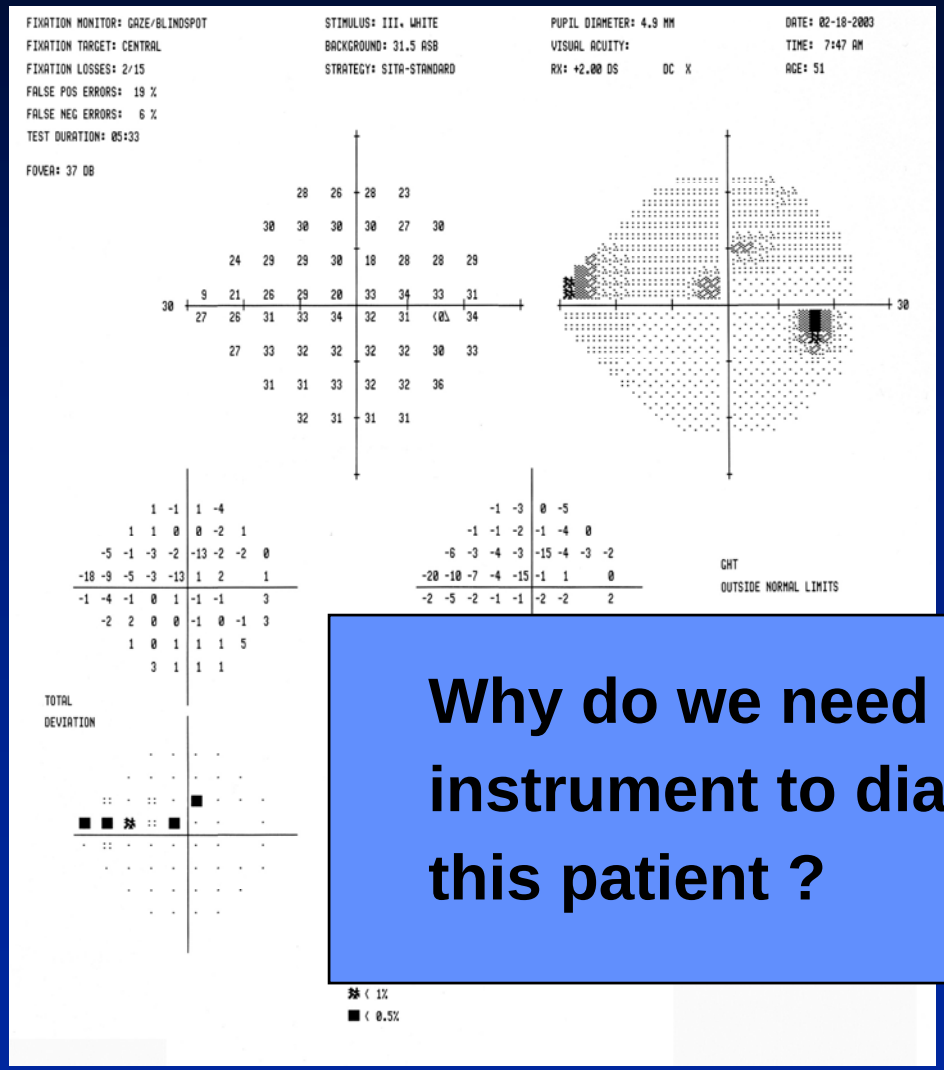
Diagnostic Studies in Imaging

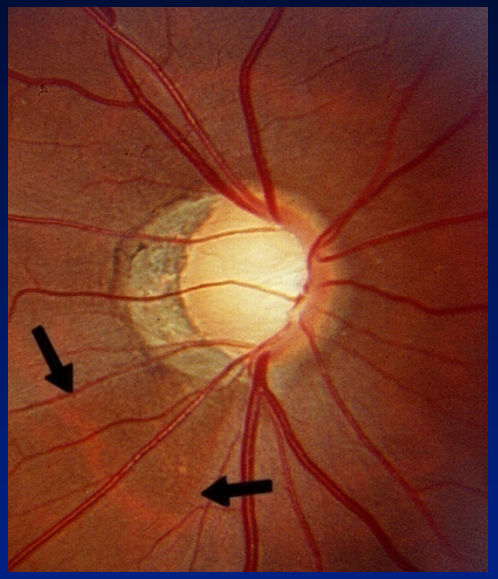
- Diagnostic Studies in Glaucoma
  - **Cases:** Glaucoma patients with **repeatable visual field loss**
  - **Controls:** Healthy individuals (volunteers without any suspicious finding)

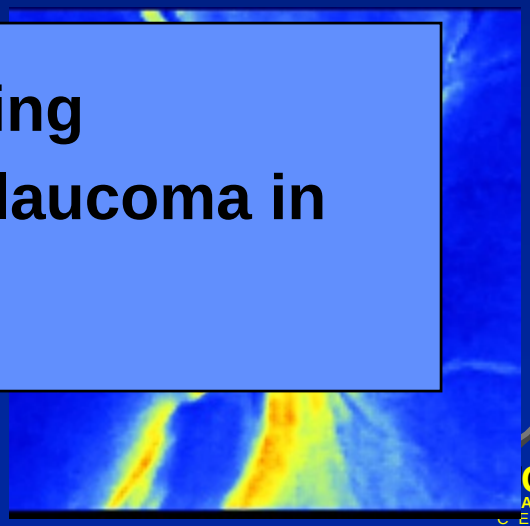
- Diagnostic accuracy measures
  - Proportion of cases correctly identified by the test as being abnormal (sensitivity)
  - Proportion of controls correctly identified by the test as being normal (specificity)
  - ROC curves

UCSD GLAUCOMA CENTER

Typical cases included in most studies…





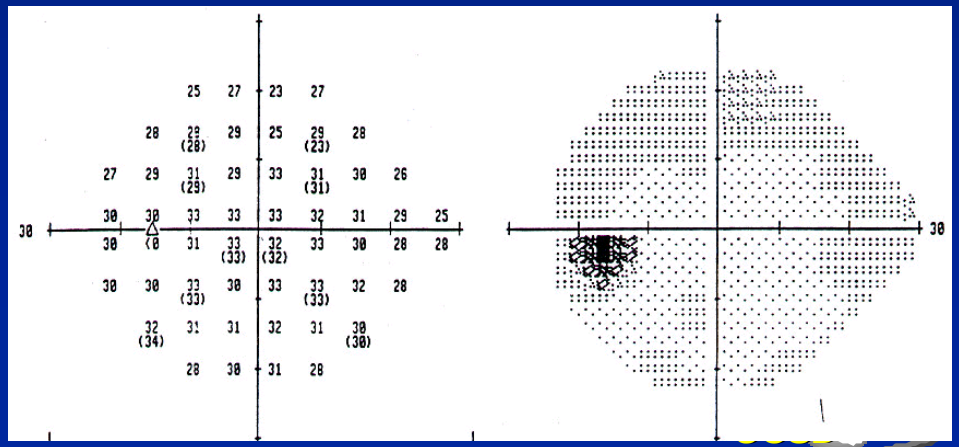


Why do we need an imaging instrument to diagnose glaucoma in this patient ?

CSD AUCOMA CENTER

A typical control… healthy volunteer

IOP = 10 mmHg



Limitations of “conventional” studies

- Case-control studies including only patients with well-defined disease and healthy subjects

  … are important for an initial evaluation of a diagnostic test

- **However...**

... In clinical practice, diagnostic tests are used to evaluate patients who are suspected of having the disease, not patients with confirmed disease

“when the diagnosis is obvious to the eye, we don’t need further diagnostic tests”

(Straus SE et al. Evidence-Based Medicine: How to Practice and Teach it)

UCSD GLAUCOMA CENTER

Limitations of “conventional” studies

- The population sample included in most studies may not be representative of the one in which we apply the diagnostic tests in everyday practice

- Estimates of sensitivity and specificity obtained from these studies may not be directly applicable in clinical practice

UCSD GLAUCOMA CENTER

EDITORIAL

How should diagnostic tests be evaluated in glaucoma?

Felipe A Medeiros

Br J Ophthalmol. 2007 Mar;91(3):273-4

Do test results distinguish patients with and without the target disorder among those in whom it is clinically sensible to suspect the disorder?

"If sensitivity is determined in seriously ill subjects and specificity in clearly healthy individuals, both will be grossly overestimated"

(Sackett)

UCSD GLAUCOMA CENTER

Heidelberg Retina Tomograph
OU Report

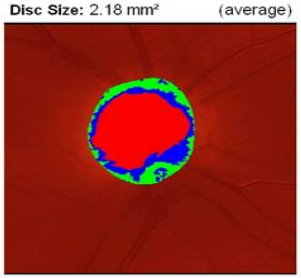
HEIDELBERG ENGINEERING

Patient: [redacted]
Pat-ID: [redacted]

DOB: May/17/1945 **Examination:** May/7/2001
Gender: female **Ethnicity:** (Caucasian)

Initial Report

OD		OS
Quality: Excellent (SD 9 µm)		**Quality:** Very good (SD 10 µm)
Focus: 1.00 dpt		**Focus:** 1.00 dpt
Operator: ---		**Operator:** ---
Disc Size: 2.18 mm² (average)		**Disc Size:** 2.16 mm² (average)



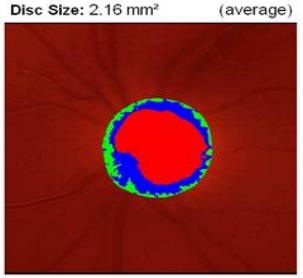
CUP

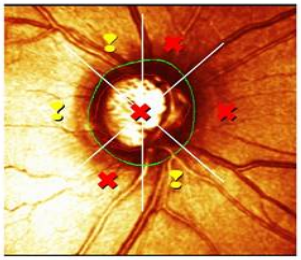
Linear Cup/Disc Ratio [ ]

0.76	Asymmetry 0.01 ✓	0.75
p = 0.03	p = 0.35	p = 0.03

Cup Shape Measure [ ]

-0.01	Asymmetry 0.02 ✓	-0.03
p = 0.002	p = 0.32	p = 0.007





MRA: Outside normal limits

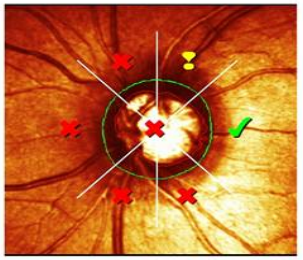
RIM

Rim Area [mm²]

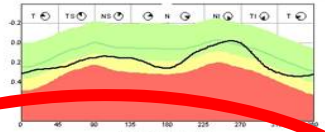
0.92 ✗	Asymmetry -0.03 ✓	0.95 ✗
p < 0.001	p = 0.42	p < 0.001

Rim Volume [mm³]

0.19	Asymmetry -0.04 ✓	0.23
p = 0.008	p = 0.26	p = 0.04



MRA: Outside normal limits

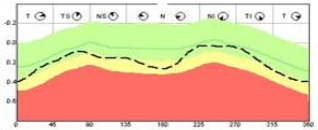


RNFL Profile

RNFL

Height Variation Contour [mm]

0.37 ✓	Asymmetry 0.00 ✓	0.37 ✓
p > 0.5	p = 0.31	p > 0.5

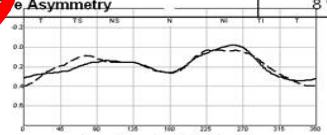


RNFL Profile

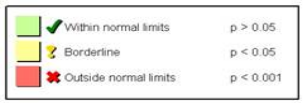
Mean RNFL Thickness [mm]

0.20 ✓	Asymmetry -0.06 ✓	0.26 ✓
p = 0.3	p = 0.09	p > 0.5

Inter-Eye Asymmetry 8 %



Combined RNFL Profile



Comments:

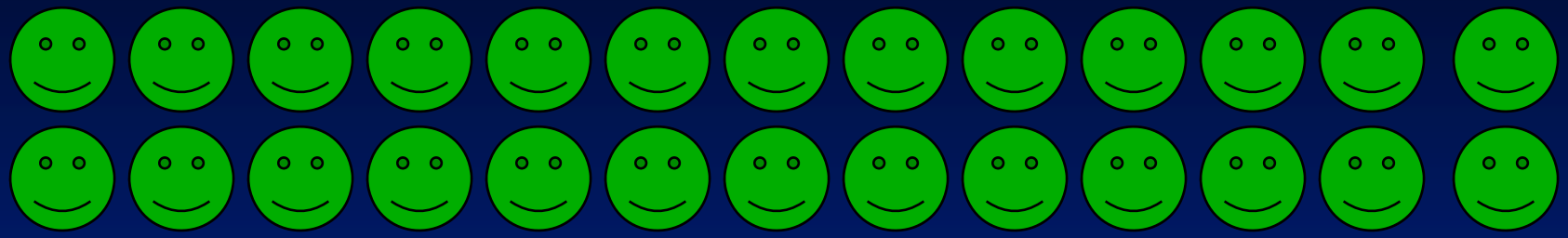
Signature:
Date: 1/28/2006

Software Version: 3.0.2.7/802
www.HeidelbergEngineering.com

UCSD GLAUCOMA CENTER

Healthy sample, recruited from the general population

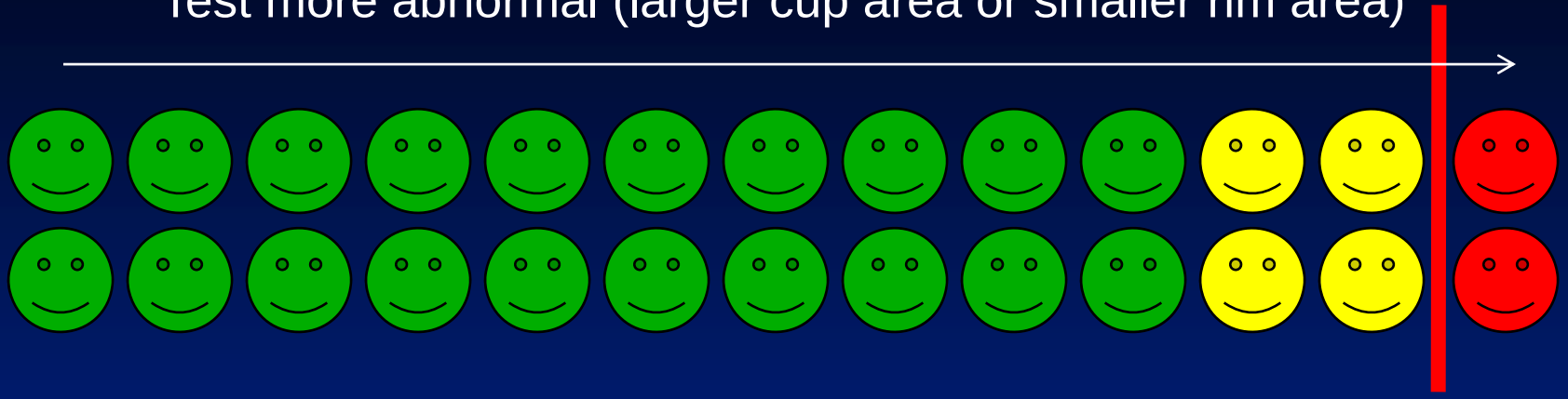
Test more abnormal (larger cup area or smaller rim area)



UCSD GLAUCOMA CENTER

Healthy sample, recruited from the general population

Test more abnormal (larger cup area or smaller rim area)



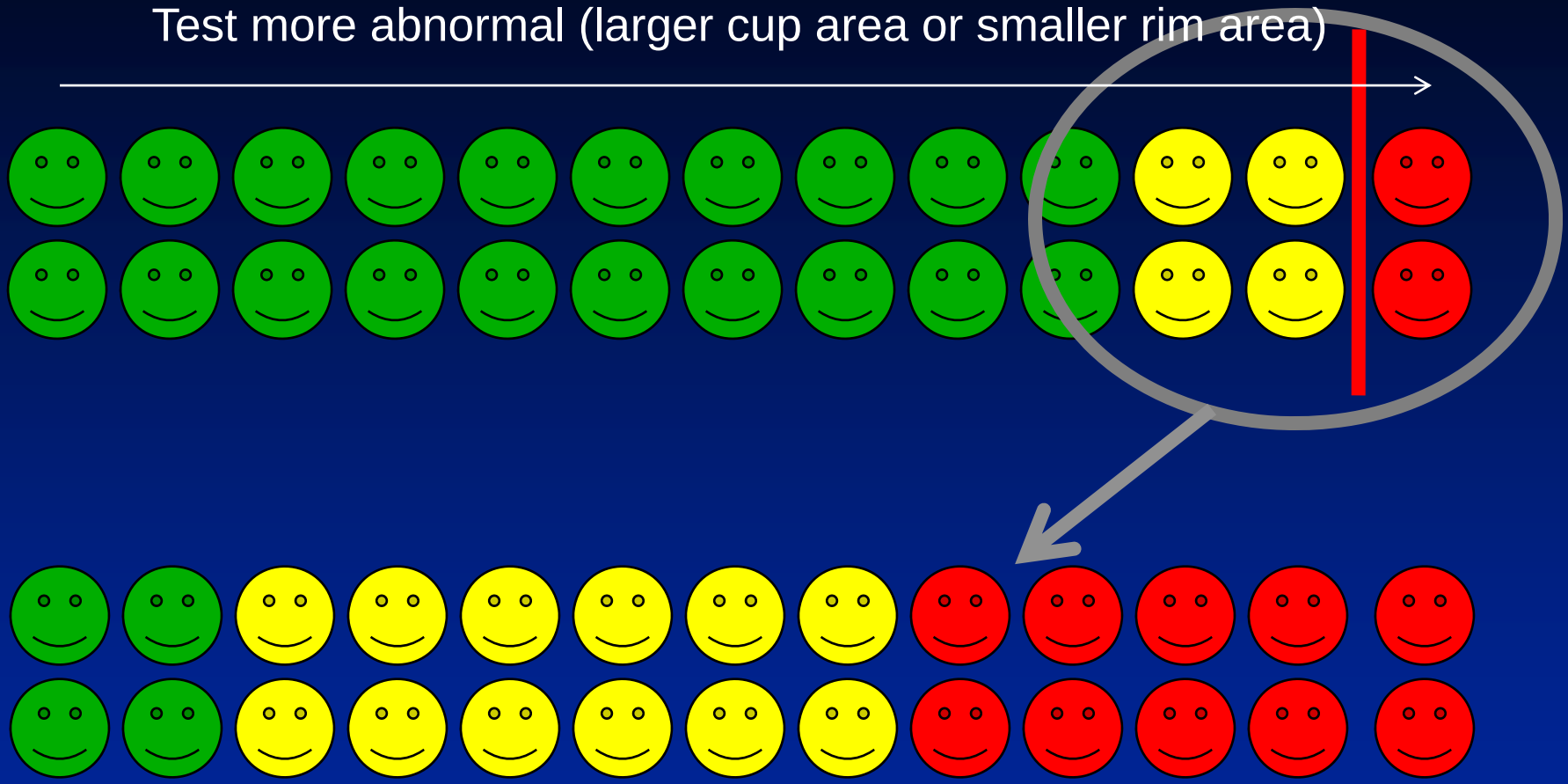
Cut-off to determine abnormality
Only 1% false-positives
For example, abnormal MRA

Should I expect the test to have only 1% false positives in my clinical practice?

UCSD GLAUCOMA CENTER

Healthy sample, recruited from the general population

Test more abnormal (larger cup area or smaller rim area)



Clinical population will be enriched by individuals with suspicious discs. The test will not have the same specificity

UCSD GLAUCOMA CENTER

If the purpose of the test is to complement current clinical evaluation…

How should we design diagnostic accuracy studies?

UCSD GLAUCOMA CENTER

An example:

Accuracy of ECG for diagnosing myocardial infarction

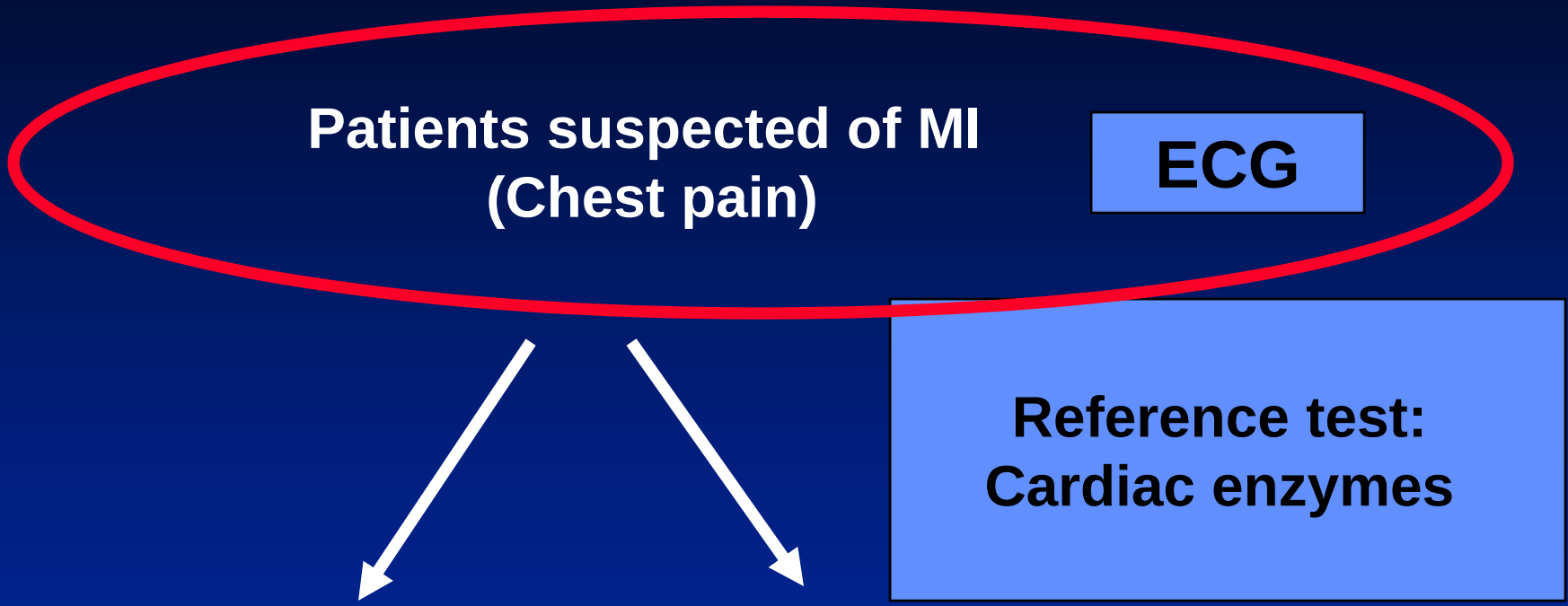
UCSD GLAUCOMA CENTER

Accuracy of ECG for myocardial infarction

- Review of diagnostic accuracy studies of ECG[1]
- All included patients were suspect of MI at the time of the ECG
- Diagnosis of MI was subsequently confirmed or ruled out based on levels of cardiac enzymes (reference test)

1. Panju AA, Hemmelgarn BR, Guyatt GH, Simel DL. The rational clinical examination. Is this patient having a myocardial infarction? JAMA 1998;280:1256-63.

Diagnostic Accuracy of ECG



1. Panju AA, Hemmelgarn BR, Guyatt GH, Simel DL. The rational clinical examination. Is this patient having a myocardial infarction? JAMA 1998;280:1256-63.

UCSD GLAUCOMA CENTER

How should we design diagnostic accuracy studies in glaucoma?

UCSD GLAUCOMA CENTER

Diagnostic Accuracy of Test X

Patients suspected of having glaucoma
(High IOP, suspicious discs, etc)

TEST X

Reference test

Glaucoma

No glaucoma

UCSD GLAUCOMA CENTER

What reference standard should be used?

Visual Field – Not a good reference standard if you want to evaluate additional clinical benefit of test X

UCSD GLAUCOMA CENTER

What reference standard should be used?

Cross-sectional **optic disc evaluation (photos, slit-lamp)**

Not a good reference standard because of poor accuracy, prognostic value

UCSD GLAUCOMA CENTER

What reference standard should be used?

Diagnostic studies involving glaucoma suspects will require longitudinal follow-up (historical or prospective)

UCSD GLAUCOMA CENTER

Is this a good design?

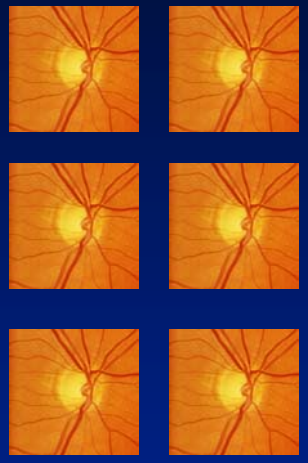
This is a prognostic study

Important to assess measures of prognostic ability (hazard ratios not enough)

Need to account for other risk factors

Test may have a weak prognostic ability but be a good diagnostic test

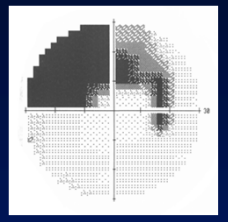
SUSPECTS



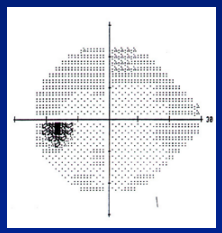
Apply the diagnostic test
BASELINE

Follow them over time
Visual fields, photos

Progress:
Glaucoma



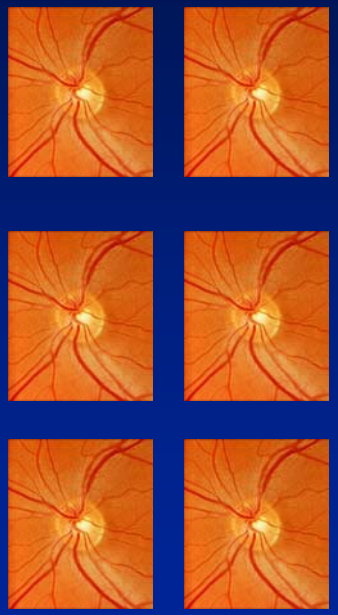
Do not progress



UCSD GLAUCOMA CENTER

Test may have a weak prognostic ability but be a good diagnostic test

Ocular hypertensive eyes
WITHOUT nerve damage



Follow them over time

Progressive nerve damage:
Glaucoma

No progression

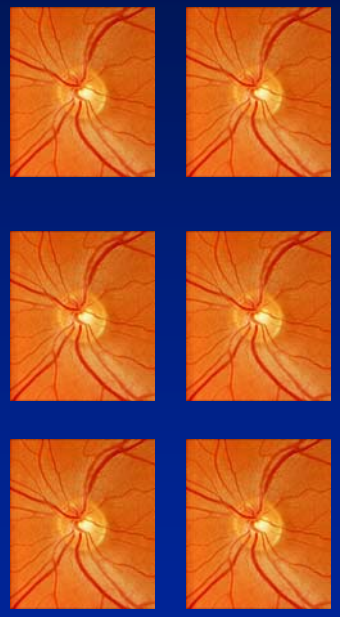
Imaging tests will not be able to discriminate these eyes at baseline (they don't have damage)
Poor prognostic value

Imaging tests may clearly separate eyes with nerve damage from those without damage here
Good diagnostic value

GLAUCOMA
CENTER

Test may have a weak prognostic ability but be a good diagnostic test

Ocular hypertensive eyes
WITHOUT nerve damage



Follow them over time

Progressive nerve damage:
Glaucoma

No progression

How can we assess diagnostic performance here if we cannot use visual fields as reference standard?
Historical follow-up

UCSD GLAUCOMA CENTER

Progressive Optic Disc Damage is Highly Predictive of Development of Functional Loss in Glaucoma

Predictor	R^2 (95% confidence interval)
Optic Disc Progression	79% (65% - 87%)
Baseline GON grading	21% (9% - 37%)
Baseline vertical C/D ratio	21% (8% - 37%)
Baseline IOP	10% (2% – 22%)
CCT	6% (1% – 15%)
Baseline PSD	26% (15% - 40%)
Age	23% (11% - 39%)

Medeiros FA, et al. Arch Ophthalmol 2009; 127: 1250-6.

UCSD GLAUCOMA CENTER

Does it matter?

UCSD
GLAUCOMA
CENTER

The Effects of Study Design and Spectrum Bias on the Evaluation of Diagnostic Accuracy of Confocal Scanning Laser Ophthalmoscopy in Glaucoma

Felipe A. Medeiros, Diana Ng, Linda M. Zangwill, Pamela A. Sample, Christopher Bowd, and Robert N. Weinreb

- **Influence of the studied population on the Diagnostic Accuracy of CSLO (HRT)**

Medeiros FA, et al. IOVS 2007; 48: 214-22.

UCSD GLAUCOMA CENTER

The Effect of Study Design…

- Compared Diagnostic Accuracy of CSLO in 2 scenarios:

- **ANALYSIS 1**
  - Discriminate patients with glaucomatous visual field loss from healthy subjects (recruited from general population)

- **ANALYSIS 2**
  - Discriminate suspects who have glaucoma from suspects who do not have glaucoma (using history of previous optic disc progression as reference standard)

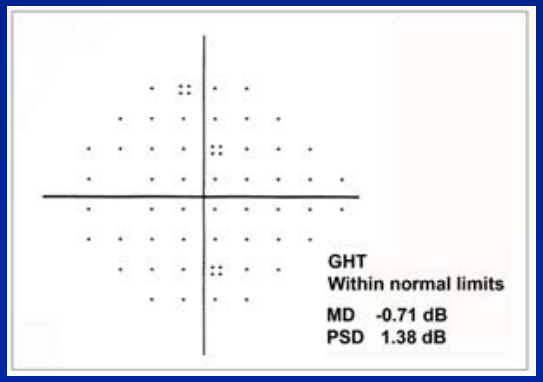
Medeiros FA, et al. IOVS 2007; 48: 214-22.

UCSD GLAUCOMA CENTER

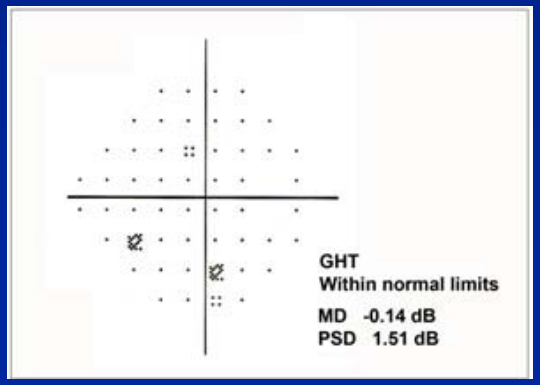
Analysis 2

Cohort of Glaucoma Suspects

Suspect A



Suspect B

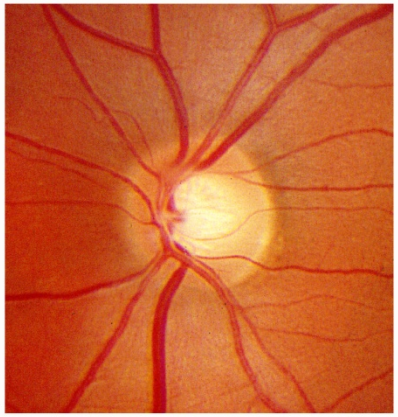


UCSD GLAUCOMA CENTER

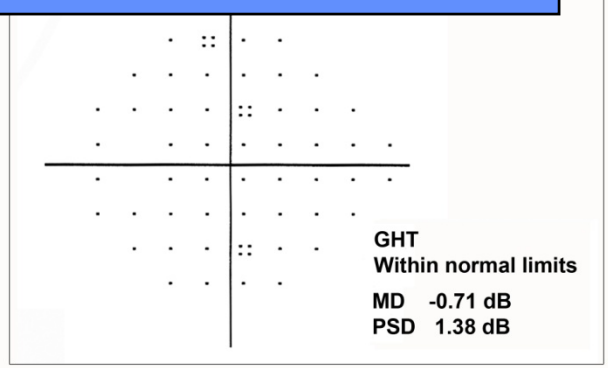
Analysis 2

Suspect, but normal

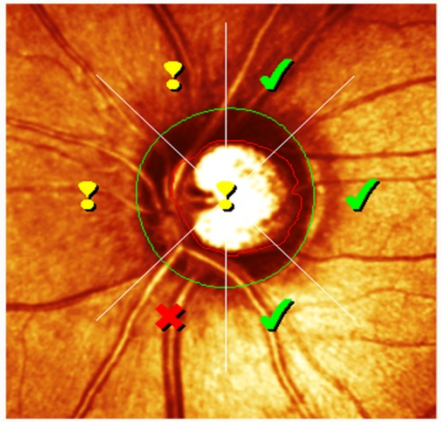
11/22/2005



A



11/17/2005



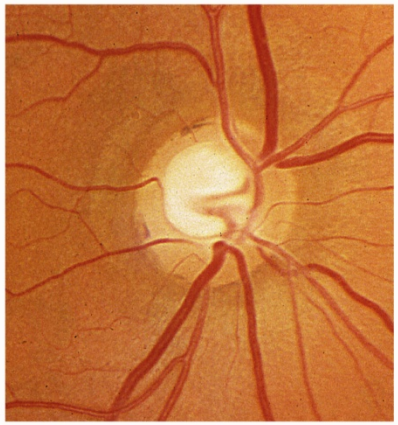
Patient followed for 18 years without treatment and without any changes to the optic nerve and VF

Analysis 2

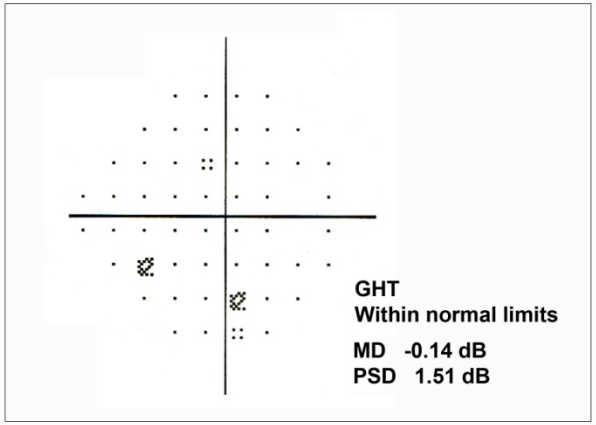
Evidence of PROGRESSIVE glaucomatous damage confirms diagnosis of glaucoma

B

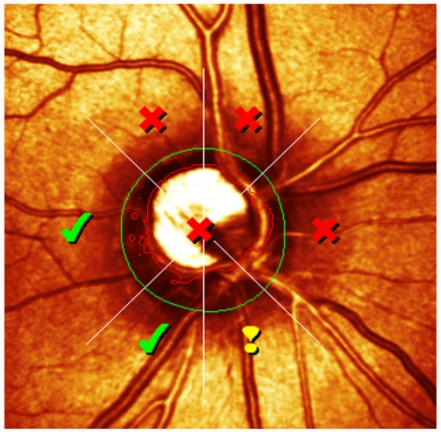
3/26/2002



3/26/2002



3/26/2002



The Effect of Study Design…

- Diferences in diagnostic accuracy in the two analyses
  - Glaucoma Probability Score (GPS)

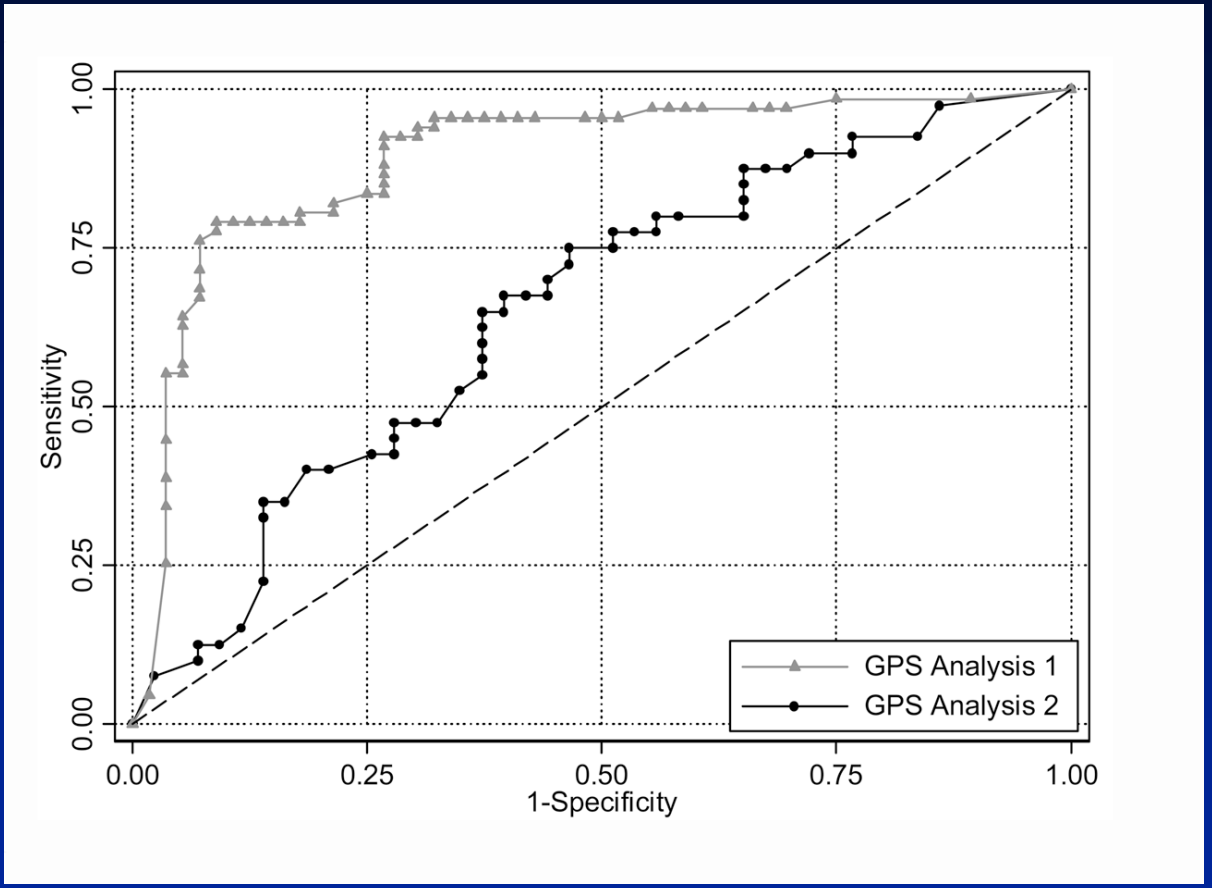
Medeiros FA, et al. IOVS 2007; 48: 214-22.

UCSD GLAUCOMA CENTER

The Effect of Study Design…

Analysis 1
ROC curve area = 0.89

Analysis 2
ROC curve area = 0.65



Medeiros FA, et al. IOVS 2007; 48: 214-22.

UCSD GLAUCOMA CENTER

Spectrum Bias

- If patients are referred as suspicious of glaucoma by the same characteristic measured by the diagnostic test (e.g., rim thinning), the test will have little additional value for decreasing diagnostic uncertainty

UCSD GLAUCOMA CENTER

Potential Applications of Imaging Instruments

- **Diagnosis**
  - **Decrease diagnostic uncertainty in those SUSPECTED of a condition (suspected of having damage or suspected of having progression)**
- **Screening**
  - **Identify abnormal or suspected cases in the general population or pre-selected subjects (e.g., older population, positive family history)**
- **Prognosis**
  - **Determine risk of developing a condition**

UCSD GLAUCOMA CENTER

Screening

- Estimates of diagnostic accuracy obtained in clearly glaucomatous versus healthy eyes seem to be more relevant to opportunistic screening situations

UCSD GLAUCOMA CENTER